Robust Graph Regularized Non-negative Matrix Factorization for Image Clustering

Xiangguang Dai[1], Keke Zhang[1], Juntang Li[2], Jiang Xiong[1], and Nian Zhang[3,*],

[1]Key Laboratory of Intelligent Information Processing and Control of Chongqing Municipal Institutions
of Higher Education, Chongqing Three Gorges University, Chongqing, 40044 China

[2]Sate Grid Chongqing Yongchuan Electric Power Supply Branch

[3]Department of Electrical and Computer Engineering University of the District of Columbia
Washington, D.C., Washington 20008 USA

Emails: daixiangguang@163.com, xk_zhang0924@126.com, 398340239@qq.com,
xjcq123@126.com, nzhang@udc.edu

Abstract—**Non-negative matrix factorization and its variants have been utilized for computer vision and machine learning, however, they fail to achieve robust factorization when the dataset is corrupted by noises. In this paper, we propose a roust graph regularized non-negative matrix factorization method (RGRNMF) for image clustering. To improve the clustering effect on the dataset corrupted by noises, we propose a weight sparse constraint on the noise matrix and impose the manifold structure into the learned parts-based representation. Experimental results demonstrate that RGRNMF can achieve the best clustering performances on the face dataset corrupted by Salt and Pepper noise and Contiguous Occlusion.**

Keywords—**Graph regularization; dimensionality reduction; non-negative matrix factorization; noise**

I. INTRODUCTION

Clustering for computer vision and subspace learning is a challenging work. Many clustering methods were proposed for image retrieval [1], image indexing [2] and image classification [3]. For image clustering effectively, a widely used approach is to discover an effective low-dimensional representation for the original data. Therefore, a lot of researches were presented to dig out the geometrical structure information of the original data, which can lead to a more discriminative representation.

In the past decades, many representation techniques were proposed including principal components analysis (PCA) [4] and non-negative matrix factorization (NMF) [5]. Among these data representation methods, the non-negative property of the learned representation leads to be more meaningful in image representation. NMF decomposes the original data matrix into two low-dimensional matrices (i.e. a basis matrix and an encoding matrix), whose product can be best approximate to the original data matrix. Due to the excellent property of NMF, some variants [6]–[19] based on NMF were proposed to improve the clustering accuracy from different views.

Although traditional NMF performs very well in learning a parts-based representation for clustering, it fails to achieve clustering while the original data is heavily corrupted. In the view of most of researchers, the loss function of traditional NMF are very sensitive to outliers. Simply, the Frobenius norm enlarges the approximation error between the original data matrix and the product of the decomposed matrices. To address this issue, some studies [6]–[12] proposed some robust loss functions to minimize the reconstruction error. These proposed methods can reduce outliers of the representation, but they cannot remove outliers. Moreover, the learned representation cannot respect the geometrical structure of the original data corrupted by noises.

To address above-mentioned problems, we present a robust graph regularized non-negative matrix factorization (RGRNMF) for image clustering. Firstly, we propose a robust weight framework to measure the approximation error. Secondly, we

construct a weight graph to encode the geometrical information of the original data corrupted by noises. Our achievements are as follows:

- We propose a robust non-negative matrix factorization framework to remove outliers, and we incorporate the geometrical information of the original data into the learned representation.
- Extensive experiments demonstrate that our proposed framework can achieve image clustering from the original data corrupted by Salt and Pepper noise or Contiguous Occlusion.

II. RELATED WORKS

Supposed that there are n sample images $\{x_i\}_{i=1}^n$ and any image x_i has $m-$dimensional features. Thus, we denote the original data matrix by $V \in R^{m \times n}$. Due to the high-dimensional property of V, it is a challenging task to achieve image clustering. Generally, NMF is utilized to find two low-dimensional matrices $W \in R^{m \times r}$ and $H \in R^{r \times n}$ for clustering. NMF can be defined by

$$V \approx WH, \tag{1}$$

where r is a factorization rank and $r \ll \min\{m, n\}$. Generally, problem (1) can be transformed into a non-convex optimization problem as follows:

$$\min_{W,H} \quad Error(V, WH)$$
$$s.t. \quad W \geq 0, H \geq 0. \tag{2}$$

where the loss function $Error$ can be the Frobenius norm, L_1 norm, $L_{2,1}$ norm or Huber. Recently, Guan et al. [12] proposed a Truncated Cauchy loss (CauchyNMF) to reduce outliers, which can be summarized as follows:

$$\min_{W \geq 0, H \geq 0} F(W, H) = \sum_{i=1}^m \sum_{j=1}^n g\left(\frac{(V - WH)_{ij}}{\gamma}\right),$$
$$\tag{3}$$

where $g(x) = \begin{cases} ln(1 + x), & 0 \leq x \leq \sigma \\ ln(1 + \sigma), & x > \sigma \end{cases}$; σ and γ denote the scale parameter and the truncation parameter. σ can be obtained by three-sigma-rule, and γ is given by the Nagy algorithm [12]. Traditional NMF utilizes the different loss functions to reduce

outliers, but they cannot remove outliers. Therefore, a robust NMF framework was proposed to eliminate outliers as follows:

$$\min_{W,H,E} \quad loss(M, WH, E) + \lambda\Omega(E, W, H)$$
$$s.t. \quad W \geq 0, H \geq 0, \tag{4}$$

where M is the original data matrix corrupted by noises, E is an error matrix, λ is a hyper-parameter, and the function Ω is the constraint term. Zhang et al. [11] proposed the Frobenius norm as the loss function and the L_1 norm as the constraint on E, which can be described as follows:

$$\min_{W,H,E} \parallel M - WH - E \parallel_F^2 + \lambda \parallel E \parallel_M$$
$$s.t. \quad W \geq 0, H \geq 0, \tag{5}$$

where $\parallel E \parallel_M = \sum_{ij} |e_{ij}|$.

III. ROBUST GRAPH REGULARIZED NON-NEGATIVE MATRIX FACTORIZATION

A. Model Formulation

Previous NMF models have some defects: 1) They cannot remove outliers from the dataset corrupted by Salt and Pepper noise or Contiguous Occlusion. 2) While the dataset is corrupted by noises, the learned representation H cannot preserve the geometrical structure information.

In (5), Zhang et al. [11] supposed that the error matrix E is sparse, but the outliers in E are neglected. If the error matrix contains some outliers, then the constraint $\parallel E \parallel_M$ is not inappropriate for outliers. Supposed that all outliers of the corrupted image matrix $M \in R^{m \times n}$ produced by Salt and Pepper noise or Contiguous Occlusion are detected. A weight graph S can be utilized to label these outliers as follows:

$$S_{ij} = \begin{cases} 0, & \text{if the pixel } M_{ij} \text{ is an outlier,} \\ 1, & \text{otherwise,} \end{cases} \tag{6}$$

Thus, we propose the constraint on E by the following form:

$$\parallel E \otimes S \parallel_M \tag{7}$$

To learn the geometrical structure information of the original data, manifold regularization is proposed to construct the relation between the original data and the low-dimensional representation. A

widely used manifold regularization term [20] can be described as follows:

$$tr(H(D-U)H^T), \qquad (8)$$

where $U_{jl} = e^{-\frac{\|x_j - x_l\|^2}{\sigma}}$ and $D_{ii} = \sum_j W_{ij}$. In summary, combining (7), (8) and (5) results in our robust graph regularized non-negative matrix factorization (RGRNMF), which can be summarized into the following optimization problem

$$\min_{W,H,E} F(W,H,E)$$
$$= \| M - WH - E \|_F^2 + \lambda \| E \otimes S \|_F^2 \qquad (9)$$
$$+ \gamma tr(H(D-U)H^T)$$
$$s.t. \quad W \geq 0, H \geq 0,$$

where λ and γ are hyper-parameters.

IV. Optimization Scheme

It is obvious that problem (9) is non-convex. Therefore, the global optimal solution cannot be searched. Supposed that the $k-$th solution of problem (9) is obtained. We can have the $k + 1-$th solution by optimizing the following problems

$$E^{k+1} = \arg\min_E \| M - W^k H^k - E \|_F^2$$
$$+ \lambda \| E \otimes S \|_F^2 \qquad (10)$$

and

$$W^{k+1} = \arg\min_W \| M - WH^k - E^{k+1} \|_F^2$$
$$s.t. \quad W \geq 0 \qquad (11)$$

and

$$H^{k+1} = \arg\min_H \| M - W^{k+1}H - E^{k+1} \|_F^2$$
$$+ \gamma tr(H(D-U)H^T)$$
$$s.t. \quad H \geq 0. \qquad (12)$$

It is easy to obtain the solution of problems (10), (11) and (12) as follows:

$$e_{ij} \leftarrow \frac{m_{ij} - (WH)_{ij}}{1 + \lambda s_{ij}}. \qquad (13)$$

$$w_{il} \leftarrow w_{il} \frac{(MH^T)_{il} - (EH^T)_{il}}{(WHH^T)_{il}}, \qquad (14)$$

$$h_{lj} \leftarrow h_{lj} \frac{(W^T M)_{lj} - (W^T E)_{lj} + \gamma HU_{lj}}{(W^T WH)_{lj} + \gamma HD_{lj}}. \qquad (15)$$
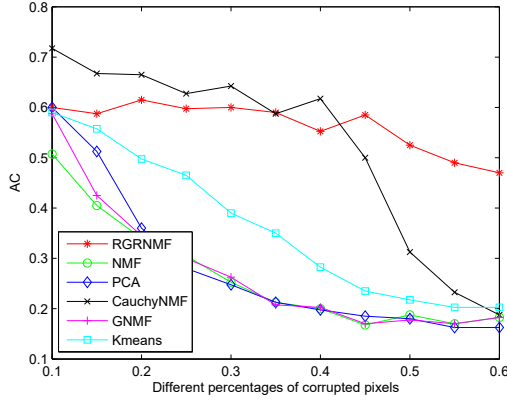
V. Experimental results

We compare our proposed method (RGRNMF) with NMF [5], RNMF [9], MahNMF [8] and CauchyNMF [12] on the clustering performances of the ORL dataset. To verify the clustering ability on the corrupted data, we propose two corruptions including Salt and Pepper noise and Contiguous Occlusion. For Salt and Pepper noise, there are several percentages of corrupted pixels from 1% to 25%. Similarly, we vary the corrupted block size for Contiguous Occlusion from 1 to 16.

To evaluate the clustering effect of all methods, we propose Accuracy (AC) and Normalized Mutual Information (NMI) [21]. Let $\lambda = 100$ and $\gamma = 100$. Figure 1 and 2 show the clustering performances on the ORL dataset contaminated by Salt and Pepper noise and Contiguous Occlusion. From these figures, we observe that:
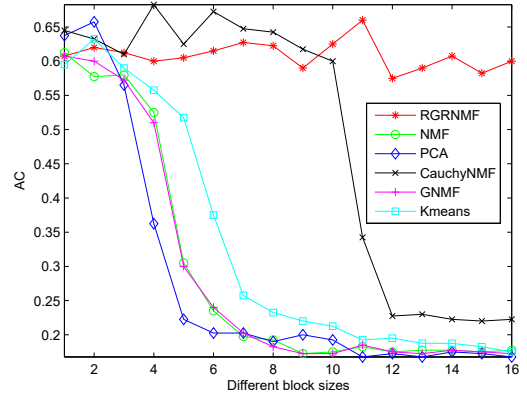
- CauthyNMF achieves satisfactory clustering ACs and NMIs from the ORL dataset corrupted by Salt and Pepper noise and Contiguous Occlusion in the beginning, however, it obtains the poor clustering effect finally. This phenomenon indicates that CauthyNMF cannot handle heavy outliers.
- NMF, PCA Kmeans and GNMF fail to achieve clustering. This means that They cannot handle outliers.
- RGRNMF has relatively stable clustering performances on the Salt and Pepper noise and Contiguous Occlusion, that is to say, RGRNMF is more robust to outliers.
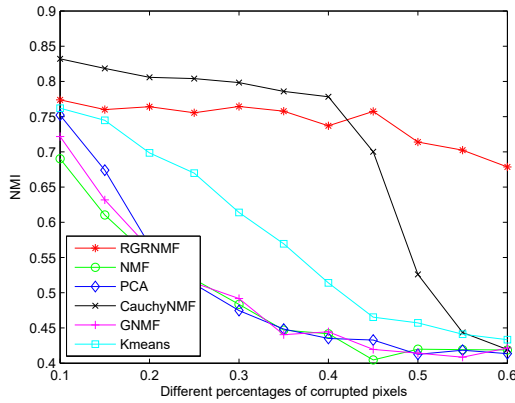
VI. Conclusion

This paper proposed robust graph regularized non-negative matrix factorization (RGRNMF) to handle Salt and Pepper noise and Contiguous Occlusion. Clustering results demonstrate that our proposed NMF framework has the following properties. Firstly, RGRNMF can learn a more effective and discriminative parts-based representation from the ORL dataset corrupted by handle Salt and Pepper noise or Contiguous Occlusion. Secondly, RGRNMF is more robust to outliers than existing NMF methods.
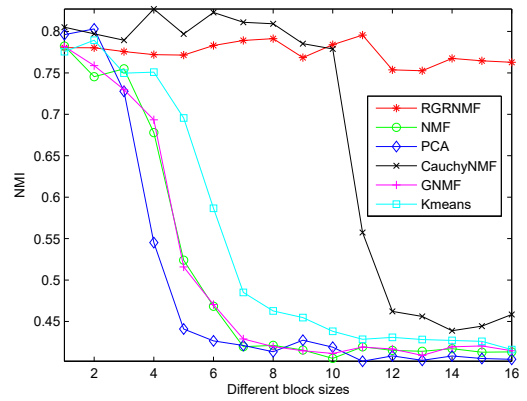
(a) Clustering AC

(b) Clustering NMI

Fig. 1: The clustering performances on the ORL dataset corrupted by Salt and Pepper noise.

(a) Clustering AC

(b) Clustering NMI

Fig. 2: The clustering performances on the ORL dataset corrupted by Contiguous Occlusion.

REFERENCES

[1] L. Chen, D. Xu, I. W. Tsang and X. Li, Spectral Embedded Hashing for Scalable Image Retrieval, *IEEE Transactions on Cybernetics*, vol. 44, no. 7, pp. 1180–1190, 2014.

[2] R. Datta, D. Joshi, L. I. Jia and J. Z. Wang, Image Retrieval: Ideas, Influences, and Trends of the New Age, *Acm Computing Surveys*, vol. 40, no. 2, pp. 35–94, 2008.

[3] B. Banerjee, F. Bovolo, A. Bhattacharya, L. Bruzzone, S. Chaudhuri and B. K. Mohan, A new self-training-based unsupervised satellite image classification technique

using cluster ensemble strategy, *IEEE Geoence and Remote Sensing Letters*, vol. 12, no. 4, pp. 741–745, 2015.

[4] M. Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, vol.3, no. 1, pp. 71–86, 1991.

[5] D. D. Lee and H. S. Seung, Learning the Parts of Objects by Non-negative Matrix Factorization, *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[6] A. B. Hamza and D. J. Brady, Reconstruction of Reflectance Spectra Using Robust Nonnegative Matrix Factorization, *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 37637–3642, 2006.

[7] D. Kong, C. Ding and H. Huang, Robust Nonnegative Matrix Factorization Using L21-norm, *In Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 673–682, 2011.

[8] N. Guan, D. Tao, Z. Luo and J. Shawetaylor, MahNMF: Manhattan Non-negative Matrix Factorization, *Journal of Machine Learning Research*, arXiv:1207.3438v1, 2012.

[9] H. Gao, F. Nie, W. Cai and H. Huang, Robust Capped Norm Nonnegative Matrix Factorization, *ACM International on Conference on Information and Knowledge Management*, pp. 871–880, 2015.

[10] L. Du, X. Li and Y. Shen, Robust Nonnegative Matrix Factorization via Half-Quadratic Minimization, *IEEE International Conference on Data Mining*, pp. 201–210, 2012.

[11] L. Zhang, Z. Chen, M. Zheng and X. He, Robust Nonnegative Matrix Factorization, *Frontiers of Electrical and Electronic Engineering in China*, vol. 6, no. 2, pp. 192–200, 2015.

[12] N. Guan, T. Liu, Y. Zhang, D. Tao and L. Davis, Truncated Cauchy Non-negative Matrix Factorization, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 246–259, 2018.

[13] G. Casalino, N. Del Buon and C. Mencar, Subtractive Clustering for Seeding Non-negative Matrix Factorizations, *Information Sciences*, vol. 257, pp. 369–387, 2014.

[14] W. Wu, Y. Jia, S. Kwong and J. Hou, Pairwise Constraint Propagation-Induced Symmetric Nonnegative Matrix Factorization, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 12, pp. 6348–6361, 2018.

[15] H. Li, K. Li, J. An, W. Zhang and K. Li, An Efficient Manifold Regularized Sparse Non-negative Matrix Factorization Model for Large-scale Recommender Systems on GPUs, *Information Sciences*, DOI: 10.1016/j.ins.2018.07.060,2018.

[16] X. Liu, W. Wang, D. He, P. Jiao, D. Jin and C. V. Cannistraci, Semi-supervised Community Detection Based on Non-negative Matrix Factorization with Node Popularity,*Information Sciences*, vol. 381, pp. 304–321, 2017.

[17] X. Peng, D. Chen and D. Xu, Hyperplane-Based Nonnegative Matrix Factorization with Label Information, *Information Sciences*, vol. 493, pp. 1–9, 2019.

[18] Z. Kang, H. Pan, S. Hoi and Z. Xu, Robust graph learning from noisy data, *IEEE transactions on cybernetics*, 2019.

[19] Z. Li, J. Tang and X. He, Robust Structured Nonnegative Matrix Factorization for Image Representation, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1947–1960, 2018.

[20] D. Cai, X. He, J. Han and T. S. Huang, Graph Regularized Nonnegative Matrix Factorization for Data Representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548-1560, 2011

[21] D. Cai, X. He and J. Han, Document Clustering Using Locality Preserving Indexing, *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 12, pp. 1624–1637, 2005.